

(The text in red is the exactly reproduced from the ISI report)

1. The EVMs used in an entire election are defined as population. Therefore, it does not guarantee the representation of each constituency. **Section 6 (page – 19): It is possible that no EVM is chosen from some Assembly segments.** It simply means that the suggested verification does not guarantee the accuracy of result for each constituency.
2. ISI report is recommending the verification of 479 EVMs in addition to the current 1 booth per constituency rule of ECI. The ISI report does not provide any statistical backing to 1 booth per constituency rule, but the report still allows ECI to continue that practice. ISI was supposed to approve or reject that practice but, surprisingly, the aspect is ignored in the analysis. **Section 6 (page –21): The practice can still be continued along with the following recommendations in this report of drawing a random sample of size 479.**
3. **The analysis presented in the report assumes that 2% defective EVMs are acceptable. Section 3.2 (page – 11): in this context of sampling of EVMs, in our view 2% or 0.02 margin of error seems appropriate.** This is a critical assumption and it is not supported by any data or argument in the report. 2% defective EVMs is very high tolerance limit. Uncertainty in 2% of the votes leads to uncertainty of the results on a huge number of constituencies. Strangely, the report does not present any sensitivity analysis for this figure. **Our recommendation portion of section 1 (page – 5): If for the each of the selected machines, the EVM count matches with the VVPAT count, then it can be concluded with an extremely high statistical confidence (more than 99.993665752% confidence) that the portion of defective EVMs is less than 2%.**
4. ECI is repeatedly and wrongly emphasizing on confidence level and ignore the margin of error used in the ISI report.
5. If any defective EVM is found in the sample size of 479, a course of action is suggested in section 5 on page 15. This procedure is just does not consider the importance of an election. All the steps suggested in ISI report can only confirm that the defective EVMs are less than 2%.
6. **Footnote on page – 4: In each of these 1521 randomly sampled EVMs, not a single mismatch of votes has been found.** This is simply wrong. The mismatches are reported in media. Although ECI claimed that the one booth/constituency rule is a confidence building exercise, the data is never made available to public on ECI website. Here are three cases of the mismatch in EVM and VVPAT counts as reported in media:
 - I. Gujarat: <https://www.news18.com/news/india/100-match-between-evms-and-paper-trail-slips-on-random-vote-count-ec-1608095.html>

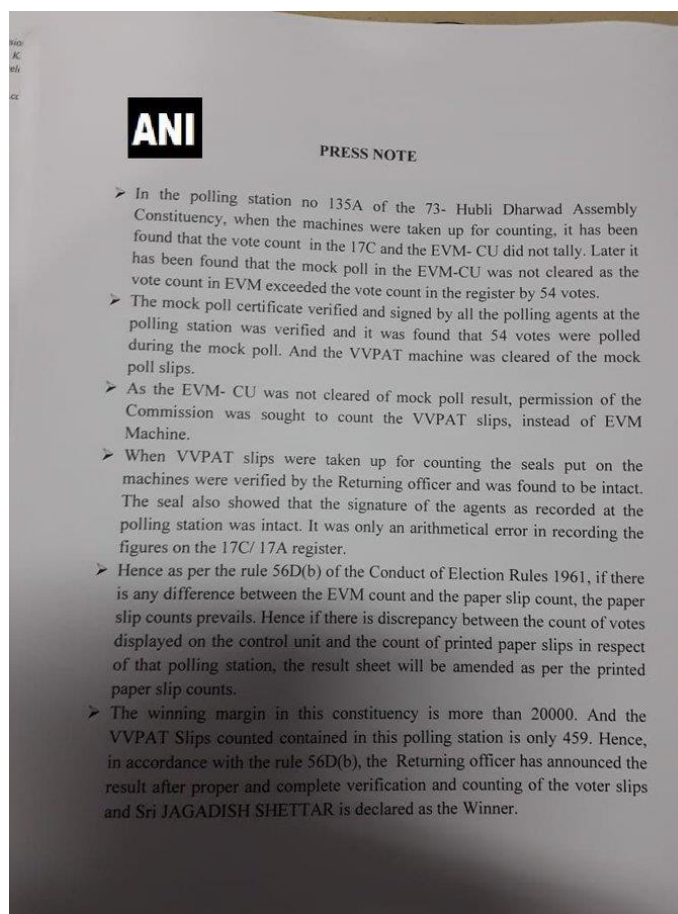
Besides, as announced earlier, VVPAT slips were taken into consideration during Monday's counting at 10 booths across seven constituencies as the presiding officers of these booths had failed to wipe out the votes from the EVMs during the mock poll ahead of the voting on December 9 and 14.

In addition, four such new cases had emerged during Monday's counting at Vagra, Dwarka, Ankleshwar and Bhavnagar- Rural seats, Swain said.

"There was a mismatch of some votes on one booth each of these four seats. This occurred because the Returning Officer must have made the same mistake but it could not be detected earlier. So we took into account VVPAT slips for these booths during the counting and resolved the issue," he said.

II. Karnataka: EC has even released a press note.

<https://twitter.com/ANI/status/996452065814528000>



III. Madhya Pradesh:

<https://www.newslaundry.com/2018/12/12/madhyapradeshlections2018-what-explains-the-delay-in-counting>

It is significant that a day before the counting, it was revealed by the State Election Commission that presiding officers at 144 polling booths in 62 Assembly seats had not cleared the mock poll data from the EVMs before the actual voting.

The matter was reported to the ECI on December 7 after which it ordered on December 10 that VVPAT slips should be counted at these 144 booths.

According to the letter of ECI, presiding officers at 144 booths had not carried out close result clear (CRC) after the mock polls which were conducted before the actual voting on 28 November. ECI issued an order that the counting at these 144 polling booths will be conducted by counting the VVPAT slips. Which would be an addition to the counting of VVPAT slips of irregularly selected machines in each and every constituency.

On the day of counting, the State Election Commission informed that the figure of such polling booths where CRC had not been removed after mock polls have increased to 175 across the state.

7. After a mismatch is found, ECI has argued (according to media reports) that the mismatch is due to the mistakes of presiding officers who forgot to erase the mock polls. But that mistake would not have been detected if VVPAT was not counted. So, the polling booths for which the VVPAT is not counted in prone to such mistakes.
8. After a mismatch is found, ECI has used the VVPAT count for those booth. But, just correcting the error in a sample does not correct the mistakes in entire population. (Analogy: a city has 1000000 children. A sample of 100 children is drawn and 5 of them are found ill. If those 5 children are treated in a hospital, it does not mean that there is no ill child in the city.) **This is important as ECI may do the same even for the larger sample size. It defeats the purpose of VVPAT verification.**
9. The standard error calculations (section 2.3 and table – 1) are just redundant for rest of the analysis. They are not used anywhere in the sample size calculation. Although, the conclusion of section 2.3 ((page – 9): **The sampling fraction n/N plays a very marginal role in the Sampling theory. For the calculation of the standard error what matters is the actual absolute value of the sample size n and not in relation to the population size N .**) will not work if EVMs used in each constituency is defined as population. For the constituency level analysis, the finite population correction (fpc) will differ significantly from 1, as tabulated here:

N	fpc	P = 0.01	P = 0.02	P = 0.05
Infinite	1	0.00445	0.00626	0.00975
600	0.4045	0.00180	0.00253	0.00393
700	0.5309	0.00236	0.00331	0.00516
800	0.6087	0.00270	0.00380	0.00591
900	0.6629	0.00293	0.00413	0.00643
1000	0.7075	0.00315	0.00443	0.00690
2000	0.8662	0.00385	0.00542	0.00844

3000	0.9130	0.00406	0.00572	0.00890
4000	0.9355	0.00416	0.00586	0.00912
5000	0.9488	0.00422	0.00594	0.00925
6000	0.9575	0.00426	0.00599	0.00933
7000	0.9637	0.00429	0.00603	0.00939
8000	0.9683	0.00431	0.00606	0.00944
9000	0.9719	0.00432	0.00608	0.00947
10000	0.9747	0.00434	0.00610	0.00950
200000	0.9988	0.00444	0.00625	0.00973

10. A better approach (My suggestion):

Number of EVMs used in a constituency should be defined as population, and 2% margin of error (allowing 2% defective EVMs) cannot be allowed. Using a hypergeometric distribution (as used in ISI report), for 99.993665752 % confidence (as used in ISI report) we will get following sample size for different population size and margin of error (MATLAB code – 1 at the end of this file):

	Margin of error (P, in percentage)							
		2.00	1.00	0.50	0.250	0.125	0.0625	0.03125
Population Size (N)	500	308	426	489	500			
	1000	380	617	854	979	1000		
	1500	410	710	1085	1385	1491		
	2000	426	763	1237	1709	1958	1999	
	2500	436	798	1344	1966	2386	2495	
	3000	443	823	1422	2171	2771	2983	3000
	3500	447	841	1483	2338	3115	3458	3500
	4000	451	855	1530	2476	3420	3916	3999
	4500	454	866	1568	2592	3692	4355	4496
	5000	457	875	1600	2690	3934	4773	4990

(Reading this table: If 2000 EVMs are used in a constituency and we want to ensure that the defective EVMs are less than 0.5%, the sample size should be 1237.)

Sample size in terms of percentage of population can be seen following table:

	Margin of error (P, in percentage)							
		2.00	1.00	0.50	0.250	0.125	0.0625	0.03125
Population Size (N)	500	61.6	85.2	97.8	100.0			
	1000	38.0	61.7	85.4	97.9	100.0		
	1500	27.3	47.3	72.3	92.3	99.4		
	2000	21.3	38.2	61.9	85.5	97.9	100.0	
	2500	17.4	31.9	53.8	78.6	95.4	99.8	
	3000	14.8	27.4	47.4	72.4	92.4	99.4	100.0
	3500	12.8	24.0	42.4	66.8	89.0	98.8	100.0
	4000	11.3	21.4	38.3	61.9	85.5	97.9	100.0
	4500	10.1	19.2	34.8	57.6	82.0	96.8	99.9
	5000	9.1	17.5	32.0	53.8	78.7	95.5	99.8

(Reading this table: If 2000 EVMs are used in a constituency and we want to ensure that the defective EVMs are less than 0.5%, the sample size should be 61.9% of population.)

- There are total 10.35 lakhs polling booths across India for a total of 543 constituencies. Average number of polling booths/constituency = 1900.
- Assuming that one EVM is used in each pooling booth, and using an average pooling booths based approach, verifying 1211 (63.7%) of VVPATs will ensure that the defective EVMs are less than 0.5%. (MATLAB code - 2 at the end of this file)
- Although, average polling booth based approach may not able to detect defective EVMs in constituencies where number of polling booths is small.
- Individual constituency based approach can also be used. The sample size for a constituency can be picked from above table based on the number of EVMs used in that particular constituency.
- If any defective EVM is found in the sample, all the EVMs of that particular constituency must be verified with the VVPAT counts.
- The VVPAT verification data should be made available for public.

MATLAB Codes:

1. To generate the data in above tables:

```
% Calculation of sample size
clc;
clear all;
P2 = 0.02;
alpha = 0.00006334248; % Corresponding to 99.993665752% confidence

for i = 1:10
    N1(i) = i*500;
    N = N1(i); % Array for population size
    for j = 1:10
        P1(j) = P2/(2^(j-1)); % Array for margin of error
        P = P1(j);
        M = P*N; % Number of defective EVMs in the population
        A = ones(10,N);
        for n = 1:N
            for k = 1:n
                % hypergeometric distribution
                % nchoosek(M,0)*nchoosek((N-M),n)/nchoosek(N,n)
                A(i,k) = (N-M-(k-1))/(N-(k-1));
            end
            B = prod(A,2);
            P0 = B(i);

            if P0<=alpha
                % Sample size for population N1(i) and margin of error P1(j)
                S(i,j) = n
                % Sample size as the percentage of population
                S1(i,j) = n*100/N;
                break;
            end
        end
    end
end
P3 = 100*P1; % Margin of error in percentage
```

2. To calculate the sample for any particular constituency:

```
% Calculation of sample size for a particular constituency
clc;
clear all;
N = 1900; % Number of EVMs used in that constituency
P = 0.5/100; % Margin of error
alpha = 0.00006334248; % Corresponding to 99.993665752% confidence
M = P*N; % Number of defective EVMs in the population

for n = 1:N
    for k = 1:n
        % hypergeometric distribution
        A(k) = (N-M-(k-1))/(N-(k-1)); % =nchoosek(M,0)*nchoosek((N-M),n)/nchoosek(N,n)
    end
    P0 = prod(A);
    if P0<=alpha
        S = n; % Sample size
        S1 = n*100/N; % Sample size as percentage of population
        break;
    end
end
end
```